Integrating Large Language Models and Möbius Group Transformations for Temporal Knowledge Graph Embedding on the Riemann Sphere

Sensen Zhang[1], Xun Liang[1]*, Simin Niu[1], Zhendong Niu[2], Bo Wu[3], Gengxin Hua[4], Long Wang[5], Zhenyu Guan[1], Hanyu Wang[1], Xuan Zhang[6], Zhiyu Li[7], Yuefeng Ma[8]

[1]School of Information, Renmin University of China
[2] Beijing Institute of Technology
[3]Xiangjiang Laboratory, Central South University
[4]Beijing Institute of Control Engineering
[5]Western Xia Research Institute, Ningxia University
[6]Harvest Fund Management Co., Ltd.
[7]Shanghai Algorithm Innovation Research Institute
[8]School of Computer, Qufu Normal University
sensen0126@ruc.edu.cn, xliang@ruc.edu.cn

Abstract

The significance of Temporal Knowledge Graphs (TKGs) in Artificial Intelligence (AI) lies in their capacity to incorporate time-dimensional information, support complex reasoning and prediction, optimize decision-making processes, enhance the accuracy of recommendation systems, promote multimodal data integration, and strengthen knowledge management and updates. This provides a robust foundation for various AI applications. To effectively learn and apply both static and dynamic temporal patterns for reasoning, a range of embedding methods and Large Language Models (LLMs) have been proposed in the literature. However, these methods often rely on a single underlying embedding space, whose geometric properties severely limit their ability to model intricate temporal patterns, such as hierarchical and ring structures. To address this limitation, this paper proposes embedding TKGs into projective geometric space and leverages LLMs technology to extract crucial temporal node information, thereby constructing the 5EL model. By embedding TKGs into projective geometric space and utilizing Möbius Group transformations, we effectively model various temporal patterns. Subsequently, LLMs technology is employed to process the trained TKGs. We adopt a parameter-efficient fine-tuning strategy to align LLMs with specific task requirements, thereby enhancing the model's ability to recognize structural information of key nodes in historical chains and enriching the representation of central entities. Experimental results on five advanced TKG datasets demonstrate that our proposed 5EL model significantly outperforms existing models.

Introduction

Knowledge graphs utilize graph models to describe knowledge and represent the association relationships between entities in the world. They play a pivotal role in large-scale data analysis, recommendation algorithms, and intelligent question-answering systems. Prominent knowledge graphs include FreeBase (Bollacker, Evans, and Paritosh 2008), WordNet (Miller 1995), ConceptNet (Speer, Chin, and Havasi 2017), and DBpedia (Lehmann, Isele, and Jakob 2015). Despite their extensive use, existing knowledge graphs often suffer from incompleteness due to limitations in data collection and the inherent complexity of the data. Consequently, essential facts may be missing. Knowledge Graph Embedding (KGE) has emerged as a key technique for addressing these gaps by leveraging representation learning to enhance and complete knowledge graphs.

In recent years, Knowledge Graph Embedding (KGE) technology has experienced rapid advancements. Researchers have developed numerous high-performing KGE models (Bordes and Usunier 2013; Sun and and 2019; Zhang et al. 2019; Cao, Xu, and Yang 2021). Additionally, many facts in the real world are inherently tied to specific time periods. For instance, Kobe was a player for the Lakers from 1996 to 2016. Temporal factors are crucial for the contextual understanding of these facts, leading researchers to introduce a temporal dimension (time t) to encapsulate the temporal context.

The Temporal Knowledge Graph (TKG) is designed to capture the dynamic nature of temporal information and real-world events. Numerous Temporal Knowledge Graph Embedding (TKGE) models (Xu, Nayyeri, and Alkhoury 2020; Sadeghian et al. 2021; Fu, Meng, and Han 2022; Chen and Wang 2022) have been developed to address the completion tasks of TKGs. However, most models focus predominantly on chain structure relations and temporal insertions, while the hierarchy and ring structure relational patterns receive less attention. Only a few relevant studies have incorporated the corresponding technical methods into constructing TKGE models. For example, HERCULES (Montella, Rojas-Barahona, and Heinecke 2021) projected TKGs into hyperbolic space, and models like DuCape (Zhang et al. 2023c) and BiQCap (Zhang et al. 2023a) combined dual quaternions and biquaternion with capsule networks to model multiplicity and hierarchical relational patterns, respectively. Despite these advancements, these methods are

*Corresponding author

limited by their underlying geometric structures, which cannot simultaneously accommodate multiple transformation types, including translation, rotation, homothety, inversion, and reflection. Consequently, most existing models fail to concurrently capture relational patterns such as chain, hierarchy, and ring structures. Furthermore, geometry-based TKGE models often struggle to leverage rich neighboring textual information and exhibit suboptimal performance in scenarios with sparse link paths.

In recent years, the advent of Large Language Models (LLMs) has led to significant advancements across various fields of research. Some researchers (Yang et al. 2024; Liu et al. 2024; Dai, Guo, and Eickhoff 2024; Zhang et al. 2024) have applied these models to the completion of temporal knowledge graphs, aiming to provide LLMs with temporally relevant logical inputs for decision-making. However, these models primarily focus on the semantic information between quadruples without adequately considering the temporal relationship patterns within the quadruples, resulting in suboptimal performance.

From the above analysis, it is evident that existing methods fall short in comprehensively modeling all relational patterns, leading to suboptimal performance. To address the intricate temporal relationship patterns in TKGs, we initially embed quadruples into projective space. By leveraging the Möbius Group transformations in projective space, we effectively model complex temporal relationships, including chains, rings, and hierarchical structures. Subsequently, the trained quadruples are processed through LLMs to extract information from neighboring nodes external to the quadruples, and a corresponding scoring function is constructed for training. Embedding into the complex underlying space of projective geometry significantly enhances the modeling capability for temporal relationship patterns such as rings and hierarchies. This method also captures the contextual information of quadruples, resulting in a more comprehensive model performance. The primary contributions of our work are as follows:

- We introduce the use of projective space transformations in the construction of TKGE models for the first time. By leveraging the five types of transformations inherent to projective space, our model effectively preserves the structure of various temporal relationship patterns within the embedding space.
- We implement a parameter-efficient fine-tuning strategy to align the LLMs with task-specific requirements, thereby facilitating the learning of crucial contextual knowledge between quadruples.
- The effectiveness of our model is validated through extensive comparisons with leading KGE and TKGE models on both link prediction and time prediction tasks across five TKG datasets, consistently outperforming established benchmarks.

Projective and Riemann Sphere

Projective

Projective geometry is a branch of mathematics that studies the properties of geometric figures that are invariant under projective transformations. Unlike Euclidean geometry, projective geometry does not involve measurements of angles and distances. Instead, it focuses on the properties that remain unchanged when figures are projected from one plane to another.

In projective geometry, the fundamental elements are points and lines, where:

- **Points and Lines**: Points are represented as sets of lines passing through a common point at infinity.
- **Homogeneous Coordinates**: Points in projective geometry are represented using homogeneous coordinates. For a point in two-dimensional space, the coordinates (x, y) are represented as $(x, y, 1)$ in homogeneous form.
- **Duality Principle**: Theorems in projective geometry remain true if points and lines are interchanged.

The complex projective line, denoted as $\mathbb{CP}^1$, is the space of lines through the origin in $\mathbb{C}^2$. Each point in $\mathbb{CP}^1$ can be represented by a pair of complex numbers $[z_0 : z_1]$, not both zero, where $[z_0 : z_1]$ and $[\lambda z_0 : \lambda z_1]$ (for $\lambda \neq 0$) represent the same point. This representation allows for a compactification of the complex plane, effectively adding a point at infinity.

Projective transformations, also known as projective mappings or homographies, are functions that map points in the projective plane to other points in the projective plane. These transformations can be represented by 3×3 matrices. A projective transformation can be represented as:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}, \tag{1}$$

where (x, y, w) are the homogeneous coordinates of the original point and (x', y', w') are the coordinates of the transformed point.

Riemann Sphere

The Riemann sphere is a model of the extended complex plane, where the complex plane is augmented with a point at infinity. It provides a natural setting for studying complex functions.

The Riemann sphere can be visualized by stereographic projection, where each point on the complex plane is projected onto the surface of a sphere. The north pole of the sphere represents the point at infinity. A complex number z in the complex plane is mapped to a point on the Riemann sphere using the transformation:

$$z \mapsto \left(\frac{2\Re(z)}{1+|z|^2}, \frac{2\Im(z)}{1+|z|^2}, \frac{|z|^2-1}{|z|^2+1} \right). \tag{2}$$

The Möbius group consists of all Möbius transformations, which are functions of the form:

$$f(z) = \frac{az+b}{cz+d}, \tag{3}$$

where a, b, c, and d are complex numbers such that $ad-bc \neq 0$.

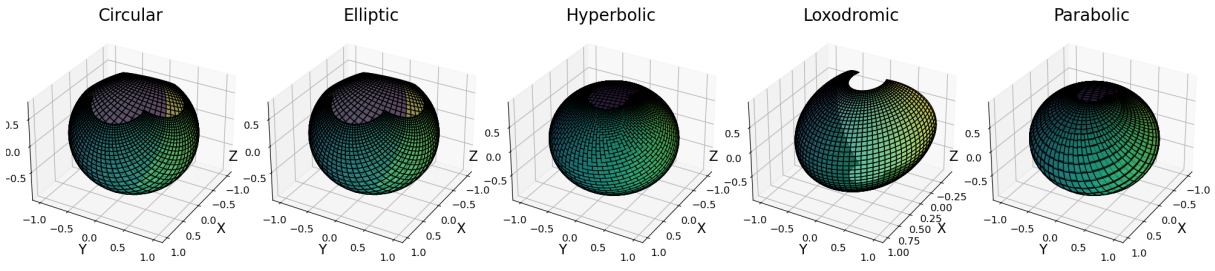

Figure 1: Transformation Functions Illustrated on a Riemann Sphere. This plot illustrates transformation functions on the complex plane, from left to right, as circular, elliptic, hyperbolic, loxodromic, and parabolic. Each subplot displays a Riemann Sphere with one or two fixed points for each transformation function.

Möbius Group

The Möbius group is isomorphic to the projective linear group $PGL(2,\mathbb{C})$. This group consists of 2×2 complex matrices with non-zero determinant, where matrices differing by a scalar multiple are considered equivalent. Specifically, a Möbius transformation:

$$f(z) = \frac{az+b}{cz+d}, \tag{4}$$

In projective geometry, the projective linear group $PGL(2,\mathbb{C})$ acts on the projective line $\mathbb{CP}^1$. Each element of $PGL(2,\mathbb{C})$ can be seen as a projective transformation of the form:

$$[z_0 : z_1] \mapsto [az_0 + bz_1 : cz_0 + dz_1], \tag{5}$$

where $[z_0 : z_1]$ are homogeneous coordinates on $\mathbb{CP}^1$.

The action of $PGL(2,\mathbb{C})$ on $\mathbb{CP}^1$ provides a powerful framework for understanding the geometric properties of Möbius transformations. These transformations can be interpreted as mappings that preserve the structure of the complex projective line, effectively describing how lines and circles are mapped to other lines and circles under projective transformations. As shown in Figure 1, the Möbius Group transformations on the complex plane can be circular, elliptic, hyperbolic, loxodromic, and parabolic. Therefore, performing Möbius Group transformations in projective space can model complex temporal relationship patterns such as cyclic, chain, and hierarchical structures. The main properties of the Möbius Group are as follows:

- **Inversion**: Möbius transformations include inversion, reflection, rotation, translation, and scaling.
- **Mapping**: These transformations map circles and lines in the complex plane to circles and lines.
- **Fixed Points**: Each Möbius transformation has one or two fixed points.

Model

TKG can be formally defined as a collection of quadruples, denoted as $\mathcal{K} = \{(h, r, o, t) \mid h, o \in \varepsilon, r \in R, t \in \boldsymbol{\tau}\}$, where a relation r exists between a head entity h and a tail entity o at time t. Here, ε represents the set of all entities, R represents the set of all relations, and $\boldsymbol{\tau}$ denotes the set of time stamps.

Entity embedding

The embeddings of the head entity h and the tail entity o are represented as $\mathbf{h} \in \mathbb{CP}^d$ and $\mathbf{o} \in \mathbb{CP}^d$, respectively, where d denotes the embedding dimension of entities. A relation r is embedded into a d-dimensional vector $\mathbf{r}$, where each element is a 2×2 matrix. $\mathbf{r}$ contains four complex vectors $\mathbf{r}_a, \mathbf{r}_b, \mathbf{r}_c$, and $\mathbf{r}_d \in \mathbb{C}^d$. Each temporal instance t is also embedded in the d-dimensional vector space, and in the model construction, we only use the time vector $\mathbf{t}$ to represent the Möbius rotation. Therefore, $\mathbf{t}$ contains four complex vectors $\mathbf{t}_a = e^{i\theta} \in \mathbb{C}^d$, with $\mathbf{t}_b = \mathbf{t}_c = 0$ and $\mathbf{t}_d = 1$.

Due to the equivalence of projective transformations on the complex projective line with transformations on the Riemann sphere, we employ transformations of the Möbius Group in our model formulation instead of transformations in the projective space. We employ a relation-specific Möbius transformation to map the head entity $\mathbf{h}$ from a source complex plane to a target complex plane $\mathbb{C}^d$. This transformation is executed through stereographic projection and subsequent application of the transformation f on and from the Riemann Sphere. To achieve this, we compute $\mathbf{h}$ to determine the element-wise transformation:

$$\mathbf{h}_r = g_r(\mathbf{h}) = f(\mathbf{h}, \mathbf{r}) = \frac{\mathbf{r}_a\mathbf{h} + \mathbf{r_b}}{\mathbf{r}_c\mathbf{h} + \mathbf{r}_d}. \tag{6}$$

Subsequently, the transformed head entity $\mathbf{h}_r$, which is specific to the relation, is translated according to the temporal sequence t as $\mathbf{h}' = g_t(\mathbf{h}_r) = \mathbf{h}_r e^{i\theta}$. The validity of triples in a Knowledge Graph (KG) is evaluated through the similarity $\langle \mathbf{h}, \mathbf{o} \rangle$ between the relation-specific transformed head $\mathbf{h}'$ and tail $\mathbf{o}$. In the context of the Möbius group, distances between points in the complex projective space $\mathbb{CP}^d$ can be computed using the Fubini-Study metric. Given two vectors $\mathbf{h}'$ and $\mathbf{o}$, the Fubini-Study distance d_{FS} between them is defined as:

$$f(h, r, o, t) = d_{\text{FS}}(\mathbf{h}', \mathbf{o}) = \arccos \sqrt{\frac{|\langle \mathbf{h}', \mathbf{o} \rangle|^2}{\langle \mathbf{h}', \mathbf{h}' \rangle \langle \mathbf{o}, \mathbf{o} \rangle}}, \tag{7}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. We use the Adam optimizer (Kingma and Ba 2015) to train our model by mini-

Datatset	#Entities	#Relations	#Training	#Validation	#Test	Time Span
ICEWS14	6869	230	72826	8941	8963	2014
ICEWS05-15	10094	251	368962	46275	46092	2005-2015
YAGO11k	10623	10	16406	2050	2051	-453-2844
Wikidata12k	12554	24	32497	4062	4062	1479-2018
GDELT	500	20	2735685	341961	341961	2015.03.31-2016.03.31

Table 1: Statistics for the various experimental datasets.

mizing the loss function (Trouillon et al. 2016) as follows:

$$L = \sum_{(h,r,o,t)\in T} \sum_{o'\in T'} log(1 + \exp(y_o f(h, r, o, t))), \quad (8)$$

where T is collections of valid quadruples and T' is generated by corrupting valid quadruples in T, and the value of y_o is defined as 1 if $(h, r, o, t) \in T$, otherwise y_o is 0.

Instruction Optimization

Instruction-tuning demonstrates exceptional zero-shot generalization capabilities by training LLMs on diverse tasks using specific instructions (Wei et al. 2022). Although fine-tuning LLMs through full-parameter updates has shown to be effective, this method faces significant scalability challenges. To address this, we build upon the Low-Rank Adaptation (LoRA) technique (Hu et al. 2022; Luo et al. 2024) by incorporating Quantized Low-Rank Adaptation (QLoRA) with adapter modules. This approach enables efficient task-specific adaptation while maintaining computational efficiency and model flexibility.

In the QLoRA approach, instead of directly using full-precision low-rank matrices, we apply quantization to the low-rank adaptation matrices W_A and W_B, further reducing memory and computational overhead. Additionally, adapter modules are used as task-specific components that sit between the layers of the LLM, facilitating more efficient and modular adaptations to different tasks. Specifically, the parameter update for the pre-trained weight matrix $W_0 \in \mathbb{R}^{d\times k}$ is defined as the product of the quantized low-rank matrices W_A and W_B, augmented by the adapter module A:

$$\delta W = \mathcal{Q}(W_A W_B) + A, \quad (9)$$

where $W_A \in \mathbb{R}^{d\times r}$ and $W_B \in \mathbb{R}^{r\times k}$ are low-rank matrices with rank $r \ll \min(d, k)$, and $\mathcal{Q}(\cdot)$ denotes the quantization operation. The adapter module $A \in \mathbb{R}^{d\times k}$ is a task-specific transformation applied after the quantized low-rank update.

The forward pass of the model, initially represented by $h = W_0 x$, is modified to incorporate the quantized low-rank update and the adapter module as follows:

$$\begin{aligned} h = W_0 x + \delta W x =& W_0 x + (\mathcal{Q}(W_A W_B) + A)x \\ =& W_0 x + (\mathcal{Q}(W_A W_B) + A)x. \end{aligned} \quad (10)$$

We employ a cross-entropy loss $\mathcal{L}$ to constrain the similarity between the predicted and ground-truth tokens, defined as:

$$\mathcal{L} = -\sum_{i=1}^{|\tilde{\mathcal{R}}|} \tilde{r}_i \log(\hat{r}_i) + (1 - \tilde{r}_i) \log(1 - \hat{r}_i), \quad (11)$$

where $\hat{\mathcal{R}}$ represents the temporal knowledge graph completion predicted by the LLMs $\mathcal{M}$, and $\tilde{\mathcal{R}}$ is the given label.

Prediction with QLoRA and Adapter Modules

The constructed instructions are fed into the trained LLMs for prediction. The response is obtained through beam search, a decoding strategy that maintains k beams of possible generated responses at each time step t. The generation process is as follows: for each generated response, the k tokens with the highest probabilities are selected based on the formula:

$$P(r_t|r_{1:t-1}) = \frac{\exp(s(r_t))}{\sum_{j=1}^{k} \exp(s(r_j))}, \quad (12)$$

where $s(r_t)$ denotes the score of the token r_t. This score can be further expanded as:

$$\begin{aligned} s(r_t) =& W_0 x_t + (\mathcal{Q}(W_A W_B) + A)x_t + b_t \\ =& W_0 x_t + (\mathcal{Q}(W_A W_B) + A + B)x_t. \end{aligned} \quad (13)$$

This results in $k \times k$ new response candidates. The next k beams of response are obtained by selecting the top k responses with the highest probabilities from the generated response candidates. The highest probability is determined by the product of probabilities of $|\hat{\mathcal{R}}|$ tokens constituting the response, where $|\hat{\mathcal{R}}|$ represents the length of the current response:

$$r_t = \arg\max_r \prod_{i=1}^{|\hat{\mathcal{R}}|} P(r_i|r_{1:i-1}). \quad (14)$$

In this context, the single-step setting is employed, wherein for each test query in the test dataset, the model can access the ground truth from past timestamps. Consequently, after the prediction for the current timestamp is completed, the ground truth from the current timestamp is added to the history of the next timestamp before its execution.

Modle	ICEWS14				ICEWS05-15			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
TransE	.280	.094	-	.637	.294	.090	-	.663
DistMult	.439	.323	-	.672	.456	.337	-	.691
ComplEx	.470	.350	.540	.710	.490	.370	.550	.730
RotatE	.418	.291	.478	.690	.304	.164	.355	.595
QuatE	.471	.353	.530	.712	.482	.370	.529	.727
TTransE	.255	.074	-	.601	.271	.084	-	.616
TA-TransE	.275	.095	-	.625	.299	.096	-	.668
HyTE	.297	.108	.416	.655	.316	.116	.445	.681
TA-DistMult	.477	.363	-	.686	.474	.346	-	.728
DE-SimplE	.526	.418	.592	.725	.513	.392	.578	.748
ATiSE	.550	.436	.629	.750	.519	.378	.606	.794
TeRo	.562	.468	.621	.732	.586	.469	.668	.795
DYERNIE	.588	.498	.638	.761	.689	.618	.728	.825
ATTH	.617	.545	.654	.754	.685	.620	.719	.806
HERCULES	.612	.543	.647	.741	.685	.621	.720	.809
ChronoR	.625	.547	.669	.773	.675	.593	.723	.820
RotateQVS	.591	.507	.642	.754	.633	.529	.709	.813
RE-GCN	.415	.309	.466	.625	.464	.352	.528	.674
TempCaps	.489	.388	.544	.679	.521	.423	.576	.705
BiQCap	.592	.563	.674	.780	.691	.612	.730	.825
GenTKG	-	.349	.480	.688	-	.360	.525	.687
HyIE	.631	.563	<u>.687</u>	<u>.786</u>	.704	<u>.625</u>	.732	.831
IME	.629	.562	.637	.781	.696	.619	.749	.831
TLogic	.430	.336	.482	.612	.470	.362	.531	.674
Vicuna	-	.315	.445	.648	-	.372	.531	.701
5HE	<u>.637</u>	<u>.569</u>	.681	.782	<u>.706</u>	.614	<u>.737</u>	<u>.837</u>
5LE	**.642**	**.571**	**.695**	**.801**	**.713**	**.632**	**.741**	**.842**

Table 2: Link prediction results on ICEWS14 and ICEWS05-15. The best results are in **bold**, and the second best results are in <u>underlined</u>.

Experiments

Dataset and Baseline

In recent years, TKGs have made significant advancements, with several benchmark datasets emerging as standards in the field. The primary datasets include ICEWS14, ICEWS05-15, GDELT, YAGO11k, and Wikidata12k. Table 1 provides a comprehensive overview of these datasets. ICEWS14 documents events from the year 2014, each annotated with precise temporal information. Similarly, ICEWS05-15 covers events spanning from 2005 to 2015, with each entry accompanied by specific timestamps. YAGO11k and Wikidata12k are derived from YAGO3, containing facts enriched with temporal annotations. GDELT, on the other hand, compiles large-scale records of human societal behaviors and events reported in news media from April 1, 2015, to March 31, 2016, incorporating various forms of temporal annotations.

We compare our models with a variety of established baselines. For static KGE models, we report the performance of TransE (Bordes and Usunier 2013), DistMult (Yang et al. 2015), ComplEx (Trouillon et al. 2016), RotatE (Sun and and 2019), and QuatE (Zhang et al. 2019). For TKGE models, we include TTransE (Leblay and Chekol 2018), TA-TransE (Niepert 2018), HyTE (Dasgupta, Ray, and Talukdar 2018), TA-DistMult (Niepert 2018), DE-SimplE (Goel and Poupart 2020), ATiSE (Xu et al. 2019), TeRo (Xu, Nayyeri, and Alkhoury 2020), DYERNIE (Han et al. 2020), ATTH (Chami et al. 2020), HERCULES (Montella, Rojas-Barahona, and Heinecke 2021), ChronoR (Sadeghian et al. 2021), TempCaps (Fu, Meng, and Han 2022), RotateQVS (Chen and Wang 2022), RE-GCN (Li et al. 2021), BiQCap (Zhang et al. 2023a), GenTKG (Liao et al. 2023), HyIE(Zhang et al. 2023b), IME(Wang et al. 2024), TLogic (Liu et al. 2022) and Vicuna (Luo et al. 2024). We have provided a detailed explanation of the aforementioned models in the section of the Appendix A. Additionally, to assess the impact of neighbor information captured by the LLMs on model performance, we developed a modified version of the original models, designated as 5HE, by excluding the LLMs component.

Evaluation Protocol

In our experiments, we utilized standard evaluation metrics for link prediction in TKGE: Mean Reciprocal Rank (MRR) and Hits at N (Hits@N). MRR is the average of the inverse ranks of correct quadruples. Hits@N measures the propor-

Modle	YAGO11k				Wikidata12k				GDELT			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
TransE	.100	.015	.138	.244	.178	.100	.192	.339	.132	.000	-	.158
DistMult	.158	.107	.161	.268	.222	.119	.238	.460	.196	.117	.208	.348
ComplEx	.167	.106	.154	.282	.233	.123	.253	.436	-	-	-	-
RotatE	.167	.103	.167	.305	.221	.116	.236	.461	-	-	-	-
QuatE	.164	.107	.148	.270	.230	.125	.243	.416	-	-	-	-
TTransE	.108	.020	.150	.251	.172	.096	184	.329	.115	.000	.160	.318
TA-TransE	.127	.027	160	.326	.178	.030	.267	.429	-	-	-	-
HyTE	.105	.015	.143	.272	.180	.098	.197	.333	.118	.000	.165	.326
TA-DistMult	.161	.103	.171	.292	.218	.122	.232	.447	.206	.124	.219	.365
ATiSE	.170	.110	.171	.288	.280	.175	.317	.481	-	-	-	-
TeRo	.187	.121	.197	.319	.299	.198	.329	.507	-	-	-	-
RotateQVS	.189	.124	.199	.323	-	-	-	-	.270	.175	.293	.458
RE-GCN	-	-	-	-	-	-	-	-	.193	.120	.206	.336
TempCaps	-	-	-	-	-	-	-	-	.258	.180	.277	.404
BiQCap	.201	.146	.226	.325	.283	.184	.314	.476	-	-	-	-
GenTKG	-	.151	.231	.378	-	.213	.261	.522	-	.184	.296	.475
HyIE	.191	.125	.201	.326	.301	.197	.328	.506	.272	.182	.292	.468
IME	-	-	-	-	-	-	-	-	.254	.187	.281	.427
Vicuna	-	.165	.241	.381	-	.228	.347	.537	-	.201	.311	.492
5HE	.194	.129	.198	.336	.304	.207	.321	.513	.264	.189	.294	.472
5LE	**.217**	**.181**	**.258**	**.397**	**.311**	**.237**	**.355**	**.546**	**.282**	**.207**	**.317**	**.501**

Table 3: Link prediction results on YAGO11k, Wikidata12k, and GDELT. Models DE-SimplE, DYERNIE, ATTH, HERCULES, and ChronoR have not experimented on datasets YAGO11k, Wikidata12k, or GDELT. The best results are in **bold**, and the second best results are in underlined.

tion of correct quadruples that appear within the top N predictions, with cut-off values of $N \in \{1, 3, 10\}$. For both MRR and Hits@N, higher values indicate better model performance. For specific implementation details of the models, please refer to the section of the Appendix B.

Experimental Results

The experiments conducted on five common TKGE datasets demonstrate that the performance of current static KGE models has significantly fallen behind that of existing TKGE-based models, highlighting the crucial importance of temporal factors in TKG link prediction. Furthermore, our constructed model achieved the best experimental results across all datasets. Notably, on the ICEWS14 and ICEWS05-15 datasets, our model exhibited substantial performance improvements over the current state-of-the-art models, as shown in Table 1. This improvement is primarily due to the more complex relational patterns within ICEWS14 and ICEWS05-15 and their shorter time spans, which allow the projective space to effectively model various temporal relationship patterns. Additionally, the stable temporal nodes enabled the effective application of the QLoRA method, enhancing the acquisition of neighbor node information and thus improving model performance.

Moreover, the experimental results reveal that our 5HE model significantly outperforms existing geometry-based TKGE models, further demonstrating the substantial impact of projective space modeling on enhancing model performance. Current deep learning-based models are now significantly outperformed by large models, primarily due to the latter's superior capability to extract deep information from key nodes and neighbor nodes within historical chains. This performance advantage is especially evident in datasets with complex relational structures. However, even within these complex datasets, our 5HE model surpasses large models and deep neural network models, indicating that these models are insufficient for link prediction in large-scale TKGs. The intricate temporal relationship patterns within TKGs play a crucial role in influencing model performance, underscoring the necessity of advanced temporal modeling capabilities.

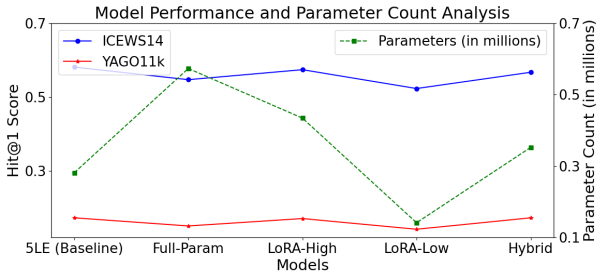

Figure 2: Ablation experiment on the ICEWS14 and YAGO11k datasets.

Ablation experiment

Impact of Geometric Rotation on Model Performance To investigate the impact of geometric rotation on model performance, we integrated existing TKGE models with our

Modle	ICEWS14				YAGO11k			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
TeRo	.562	.468	.621	.732	.187	.121	.197	.319
RotateQVS	.591	.507	.642	.754	.189	.124	.199	.323
BiQCap	.592	.563	.674	.780	.201	.146	.226	.325
HyIE	.631	.563	.687	.786	.191	.125	.201	.326
IME	.629	.562	.637	.781	-	-	-	-
TERoL	.564	.472	.627	.739	.204	.168	.245	.382
RotateQVSL	.599	.512	.647	.758	.206	.171	.252	.386
BiQCapL	.599	.568	.681	.785	.212	.174	.251	.389
IMEL	.635	.569	.642	.785	.212	.175	.255	.391
HyIEL	.638	.564	.688	.793	.215	.179	.254	.395
5LE	**.642**	**.571**	**.695**	**.801**	**.217**	**.181**	**.258**	**.397**

Table 4: Link prediction results on ICEWS14 and YAGO11k. The best results are in **bold**, and the second best results are in underlined.

constructed LLMs. In this process, geometric space rotations are captured using LLMs to extract domain-specific information. We selected models from different vector spaces: TERo from the complex vector space, RotateQVS from the high-dimensional vector space, BiQCap from the dual quaternion space, IME from Heterogeneous Geometric Subspaces, and HyIE from the hybrid vector space. These models were combined with LLMs to create TERoL, RotateQVSL, BiQCapL, IMEL, and HyIEL respectively. Experiments were conducted on the ICEWS14 and YAGO11k datasets. The experimental results are shown in Table 4.

From Table 4, it is evident that the integration with LLMs improved the performance of all models. Notably, the models showed significant performance enhancements on the YAGO11k dataset. This is primarily due to the limited temporal relationship patterns in the original models, allowing LLMs to play a more significant role. However, the performance of these models is still slightly lower than our constructed 5LE model, indicating that the projective space currently offers the best performance for modeling temporal relationship patterns.

Impact of Model Size on Model Performance To investigate how different parameter adaptation methods impact model performance, we designed an ablation experiment on the ICEWS14 and YAGO11k datasets. Our objective is to demonstrate the challenges of full-parameter updates and the effectiveness of the QLoRA method combined with the plugin encapsulation strategy of PEFT. The baseline model, 5LE, uses QLoRA with PEFT. We compared this with several ablation models: a full-parameter update model (Full-Param), models using QLoRA with varying degrees of rank reduction (LoRA-High and LoRA-Low), and a hybrid model combining full-parameter updates with LoRA and PEFT (Hybrid). The Full-Param model updates all parameters during fine-tuning, while the QLoRA models adjust the rank reduction to test its impact. The Hybrid model combines full-parameter updates with QLoRA and PEFT to leverage the strengths of both approaches.

The experiment results are shown in Figure 2. Analyzing the results, we observe that the baseline model (5LE), which uses QLoRA with PEFT, performs the best on both datasets, achieving the highest Hit@1 scores. The Full-Param model, which uses full-parameter updates, shows a noticeable drop in performance, indicating the challenges associated with this approach, especially at a large scale. The LoRA-High model, with higher rank adaptation, performs better than the Full-Param model but not as well as the baseline, suggesting that while increasing rank can help, it doesn't fully mitigate the complexity of full-parameter updates. The LoRA-Low model, with lower rank adaptation, performs the worst, highlighting the limitation of too much rank reduction. Finally, the Hybrid model, which combines full-parameter updates with QLoRA and PEFT, performs better than the Full-Param model but still falls short of the baseline, indicating that the combination approach does provide some benefits but is not as effective as using QLoRA and PEFT alone. The results demonstrate that the Low-Rank Adaptation (QLoRA) method combined with the plugin encapsulation strategy of PEFT provides an effective and efficient way to adapt large language models, maintaining high performance with fewer parameters. This approach outperforms full-parameter updates and other rank adaptation strategies, highlighting its suitability for large-scale applications.

Conclusion

To address the limitations of existing TKGE models, such as insufficient modeling capabilities, inability to leverage rich neighboring textual information, and poor performance with sparse link paths, this study proposes embedding TKGs into projective space. This approach utilizes transformations from the Möbius Group in projective space to model the complex temporal relationship patterns of TKGs. Additionally, LLMs are employed to capture contextual information of neighboring nodes, exploring undiscovered connections in the TKGs based on historical facts. Experimental results on five common TKG datasets demonstrate that our model achieves superior performance compared to existing TKGE models, validating the effectiveness of the proposed approach.

Acknowledgements

This work was supported by National Natural Science Foundation of China 62072463, Supported by the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China 24XNKJ31, and Opening Project of State Key Laboratory of Digital Publishing Technology of Founder Group. Xun Liang is the corresponding author of this paper.

References

Bollacker, K. D.; Evans, C.; and Paritosh, P. K. 2008. Freebase: a collaboratively created graph database for structuring human knowledge. In *ACM*, 1247–1250.

Bordes, A.; and Usunier, N. 2013. Translating Embeddings for Modeling Multi-relational Data. In *NIPS*, 2787–2795.

Cao, Z.; Xu, Q.; and Yang, Z. 2021. Dual Quaternion Knowledge Graph Embeddings. In *AAAI*, 6894–6902. AAAI Press.

Chami, I.; Wolf, A.; Juan, D.-C.; Sala, F.; Ravi, S.; and Ré, C. 2020. Low-dimensional hyperbolic knowledge graph embeddings. *arXiv preprint arXiv:2005.00545*.

Chen, K.; and Wang, Y. 2022. RotateQVS: Representing Temporal Information as Rotations in Quaternion Vector Space for Temporal Knowledge Graph Completion. In *ACL*, 5843–5857.

Dai, Y.; Guo, W.; and Eickhoff, C. 2024. Wasserstein adversarial learning based temporal knowledge graph embedding. *Inf. Sci.*, 659: 120061.

Dasgupta, S. S.; Ray, S. N.; and Talukdar, P. P. 2018. HyTE: Hyperplane-based Temporally aware Knowledge Graph Embedding. In Riloff, E.; Chiang, D.; Hockenmaier, J.; and Tsujii, J., eds., *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, 2001–2011. Association for Computational Linguistics.

Fu, G.; Meng, Z.; and Han, Z. 2022. TempCaps: A Capsule Network-based Embedding Model for Temporal Knowledge Graph Completion. In *Proceedings of the Sixth Workshop on Structured Prediction for NLP*, 22–31. Association for Computational Linguistics.

Goel, R.; and Poupart, S. M. K. P. 2020. Diachronic Embedding for Temporal Knowledge Graph Completion. In *AAAI*, 3988–3995. AAAI Press.

Han, Z.; Ma, Y.; Chen, P.; and Tresp, V. 2020. Dyernie: Dynamic evolution of riemannian manifold embeddings for temporal knowledge graph completion. *arXiv preprint arXiv:2011.03984*.

Hu, E. J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; and Chen, W. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In Bengio, Y.; and LeCun, Y., eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Leblay, J.; and Chekol, M. W. 2018. Deriving Validity Time in Knowledge Graph. In *Companion of the The Web Conference 2018 on The Web Conference 2018, WWW 2018, Lyon , France, April 23-27, 2018*, 1771–1776. ACM.

Lehmann, J.; Isele, R.; and Jakob, M. 2015. DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, 6(2): 167–195.

Li, Z.; Jin, X.; Li, W.; Guan, S.; Guo, J.; Shen, H.; Wang, Y.; and Cheng, X. 2021. Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning. In Diaz, F.; Shah, C.; Suel, T.; Castells, P.; Jones, R.; and Sakai, T., eds., *SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021*, 408–417. ACM.

Liao, R.; Jia, X.; Ma, Y.; and Tresp, V. 2023. GenTKG: Generative Forecasting on Temporal Knowledge Graph. *CoRR*, abs/2310.07793.

Liu, R.; Yin, G.; Liu, Z.; and Tian, Y. 2024. Reinforcement learning with time intervals for temporal knowledge graph reasoning. *Inf. Syst.*, 120: 102292.

Liu, Y.; Ma, Y.; Hildebrandt, M.; Joblin, M.; and Tresp, V. 2022. TLogic: Temporal Logical Rules for Explainable Link Forecasting on Temporal Knowledge Graphs. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, 4120–4127. AAAI Press.

Luo, R.; Gu, T.; Li, H.; Li, J.; Lin, Z.; Li, J.; and Yang, Y. 2024. Chain of History: Learning and Forecasting with LLMs for Temporal Knowledge Graph Completion. *CoRR*, abs/2401.06072.

Miller, G. A. 1995. WordNet: A Lexical Database for English. *ACM*, 38(11): 39–41.

Montella, S.; Rojas-Barahona, L. M.; and Heinecke, J. 2021. Hyperbolic Temporal Knowledge Graph Embeddings with Relational and Time Curvatures. In *ACL/IJCNLP*, 3296–3308.

Niepert, A. G. M. 2018. Learning Sequence Encoders for Temporal Knowledge Graph Completion. In *Conference on Empirical Methods in Natural Language Processing*, 4816–4821.

Sadeghian, A.; Armandpour, M.; Colas, A.; and Wang, D. Z. 2021. ChronoR: Rotation Based Temporal Knowledge Graph Embedding. In *AAAI*, 6471–6479.

Speer, R.; Chin, J.; and Havasi, C. 2017. ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In *AAAI*, 4444–4451. AAAI Press.

Sun, Z.; and and, Z. D. 2019. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. In *ICLR*.

Trouillon, T.; Welbl, J.; Riedel, S.; Gaussier, É.; and Bouchard, G. 2016. Complex Embeddings for Simple Link Prediction. In *ICML*, volume 48 of *JMLR Workshop and Conference Proceedings*, 2071–2080. JMLR.org.

Wang, J.; Cui, Z.; Wang, B.; Pan, S.; Gao, J.; Yin, B.; and Gao, W. 2024. IME: Integrating Multi-curvature Shared and Specific Embedding for Temporal Knowledge Graph Completion. In *Proceedings of the ACM on Web Conference 2024*, 1954–1962.

Wei, J.; Bosma, M.; Zhao, V. Y.; Guu, K.; Yu, A. W.; Lester, B.; Du, N.; Dai, A. M.; and Le, Q. V. 2022. Finetuned Language Models are Zero-Shot Learners. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Xu, C.; Nayyeri, M.; and Alkhoury, F. 2020. TeRo: A Time-aware Knowledge Graph Embedding via Temporal Rotation. In *COLING*, 1583–1593.

Xu, C.; Nayyeri, M.; Alkhoury, F.; and Lehmann, J. 2019. Temporal Knowledge Graph Embedding Model based on Additive Time Series Decomposition. *CoRR*.

Yang, B.; Yih, W.; He, X.; Gao, J.; and Deng, L. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In Bengio, Y.; and LeCun, Y., eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Yang, J.; Ying, X.; Shi, Y.; and Xing, B. 2024. Tensor decompositions for temporal knowledge graph completion with time perspective. *Expert Syst. Appl.*, 237(Part A): 121267.

Zhang, F.; Chen, H.; Shi, Y.; Cheng, J.; and Lin, J. 2024. Joint framework for tensor decomposition-based temporal knowledge graph completion. *Inf. Sci.*, 654: 119853.

Zhang, S.; Liang, X.; Li, Z.; Feng, J.; Zheng, X.; and Wu, B. 2023a. BiQCap: A Biquaternion and Capsule Network-Based Embedding Model for Temporal Knowledge Graph Completion. In Wang, X.; Sapino, M. L.; Han, W.; Abbadi, A. E.; Dobbie, G.; Feng, Z.; Shao, Y.; and Yin, H., eds., *Database Systems for Advanced Applications - 28th International Conference, DASFAA 2023, Tianjin, China, April 17-20, 2023, Proceedings, Part II*, volume 13944 of *Lecture Notes in Computer Science*, 673–688. Springer.

Zhang, S.; Liang, X.; Tang, H.; and Guan, Z. 2023b. Hybrid Interaction Temporal Knowledge Graph Embedding Based on Householder Transformations. In *Proceedings of the 31st ACM International Conference on Multimedia*, 8954–8962.

Zhang, S.; Liang, X.; Tang, H.; Zheng, X.; Zhang, A. X.; and Ma, Y. 2023c. DuCape: Dual Quaternion and Capsule Network-Based Temporal Knowledge Graph Embedding. *ACM Trans. Knowl. Discov. Data*, 17(7): 104:1–104:19.

Zhang, S.; Tay, Y.; Yao, L.; and Liu, Q. 2019. Quaternion Knowledge Graph Embeddings. In Wallach, H. M.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E. B.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, 2731–2741.